Estimation of β-structure content of proteins by means of deconvolved FTIR spectra

Summary

Fourier self-deconvolution was applied to the infrared spectra of five globular proteins with a high β-structure content and to the essentially α-helical protein hemoglobin. The featureless amide I′ bands around 1650 cm^{-1} were thereby resolved into six to nine components, depending on the protein. Specific components were assigned to the β-structure segments in each protein. The frequencies and the number of 'β-bands' differ from one protein to another. The areas of the components were evaluated by means of a Gauss–Newton iteration procedure. It appears that the total area of the β-bands, as a fraction of the total amide I′ band area, reflects the relative β-structure content of each protein studied.

Key words: infrared; Fourier transform; protein conformation; deconvolution.

Introduction

Fourier transform infrared (FTIR) instrumentation and software have in recent years revitalized the whole field of vibrational spectroscopy [1,2]. The development of second derivative spectroscopy [3–5] and Fourier self-deconvolution techniques [6,7] are particularly noteworthy. Qualitative application of these methods to problems of protein structure has recently been described [5,8,9]. The present communication deals with quantitative estimation of the β-structure content of several proteins by a combination of Fourier self-deconvolution and band fitting by the Gauss–Newton iterative technique.

Materials and Methods

Bovine α-chymotrypsin (Cat. No. C4129), α-chymotrypsinogen A (C4879), hemoglobin (H2500), ribonuclease A (R4875), ribonuclease S (R6000), and concanavalin

A (Con A) from *Canavalia ensiformis* (C2010) were supplied by the Sigma Chemical Company, St. Louis, MO[a]. The samples were prepared and FTIR spectra of D_2O solutions in 0.075 mm CaF_2 cells were obtained as previously described [5]. The protein concentrations were about 5% w/v. Deconvolution was carried out by the method described by Kauppinen et al. [6], as adapted for the Nicolet 7199 FTIR spectrometer. This technique requires two constants as input: s, the half width at half height of the chosen line shape function, and K, a resolution enhancement factor, which reflects the extent of narrowing of the unresolved component bands, as defined in ref. 6. (In the Nicolet software $2s$, designed as VF0, is used instead of s and K is called VF1). The proper choice of values for s and K is of paramount importance for meaningful results [6,7]. The value of s should ideally equal the inherent width of the individual unresolved components. Deconvolution increases with increasing s, but too large a value leads to 'overdeconvolution', side lobes, and grossly distorted spectra. The maximum value of K is limited by the nominal instrument resolution $\Delta\tilde{\nu}$ (2 cm^{-1} in the present work) and the signal-to-noise ratio, S/N. K should never exceed $\log(S/N)$ [6] and in any case should never be greater than about $1.5(s/\Delta\tilde{\nu})$ [7]. We have obtained good results with $s = 6.5$ cm^{-1} and $K = 2.4$. For spectra with very sharp components, values of 5 cm^{-1} and 2.0 were also used. In addition to decreasing the widths of band components, deconvolution simultaneously increases their peak heights. Thus, quantitative measurements by means of Beer's Law plots based on peak height measurements at fixed frequencies are not valid for deconvolved spectra. Fortunately, the integrated areas of component bands remain, in principle, unaffected by deconvolution [6]. The concept of 'area absorptivity' (band area per unit molar concentration and unit path length) therefore becomes important for quantitative studies.

Results and Discussion

Fig. 1 presents original and deconvolved spectra of ribonuclease A and α-chymotrypsin in the amide I′ region, Figs. 2 and 3 give the deconvolved spectra of ribonuclease S, α-chymotrypsinogen, Con A and hemoglobin. The original amide I′ bands in Fig. 1 show no fine structure, nor do the original spectra of the other four proteins. It is evident, particularly in the deconvolved spectra, that all proteins with a high β-content have the strongest component at 1630–1640 cm^{-1}, while α-helical protein hemoglobin has the dominant component centered at 1652 cm^{-1}, in agreement with well established correlations [10]. The strong β-component around 1630 cm^{-1} (the $\mathbf{B}_1$ amide I′ mode in the single chain approximation [10,11]), however, varies in frequency from one protein to another and is frequently split into several sub-components [5,9]. This is particularly striking for Con A where three

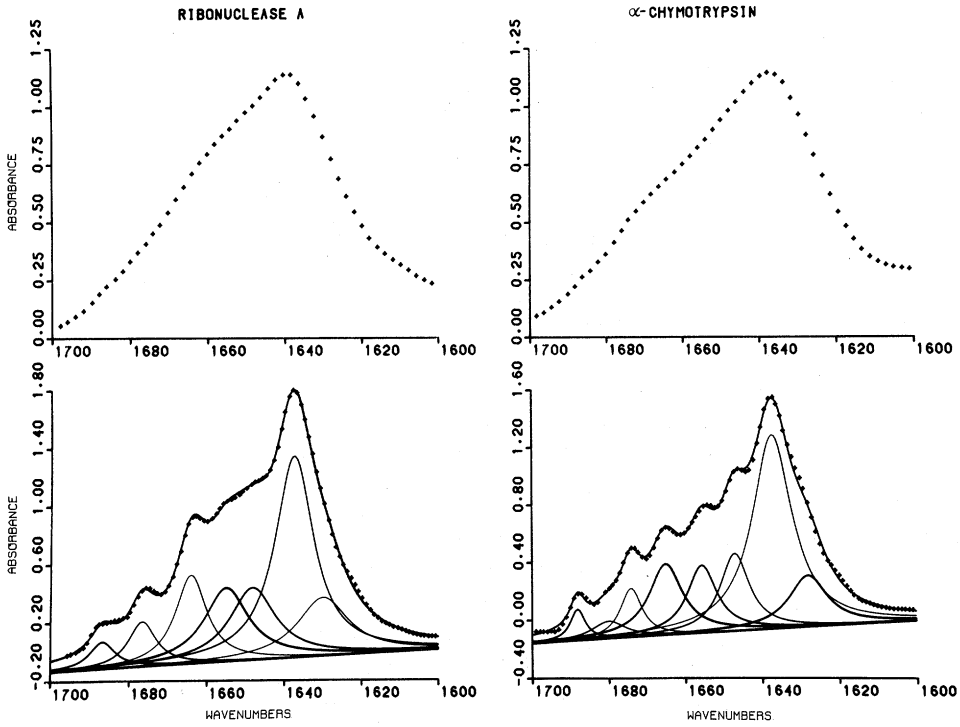

Fig. 1. Original (upper) and deconvolved (lower) FTIR spectra of ribonuclease A and α-chymotrypsin. The deconvolved spectra were fit with the smallest number of component bands needed to minimize the deviation between the calculated and experimental curves. In general, the band centers were fixed to the frequencies observed in the deconvolved spectra (as verified by second derivative data) and then the band heights and widths were allowed to attain the best values by iteration. Crosses (+ + + +) symbolize the digitized experimental data; solid lines (———) represent the individual calculated component bands and the overall calculated spectra obtained by taking the sum of their intensities. Ca. 5% w/v in D_2O with 0.01 M NaCl; pD = ca. 7. Pathlength: 0.075 mm. $s = 6.5$ cm^{-1}. $K = 2.4$.

distinct bands are observed between 1 622 and 1 639 cm^{-1} (Fig. 3) and also happens in many other proteins [8,9]. The weak $\mathbf{A}_1$ component at 1 670–1 676 cm^{-1} seems more stable. Information regarding the components associated with β-segments is summarized in Table I. The other components are not as easily assigned. Each of the β-rich proteins has bands around $1\,654 \pm 2$ cm^{-1} which could arise from helical segments [5,9,10]. It thus appears that the major band frequencies of the α-helices might also vary from one protein to another. The remaining components are associated with turns [5].

Table I also provides information about the areas of spectral components associated with β-segments (A_i), hereafter called simply 'β-bands'. For each protein, the sum of the β-band areas, as a fraction of the total amide I' (A_{tot}) area, is designed $\%A$. There is no compelling reason why $\%A$ should reflect the 'β-content' of a given protein, because the 'area absorptivities' for the various components are not necessarily equal. Nevertheless, there seems to be a correlation between $\%A$ and

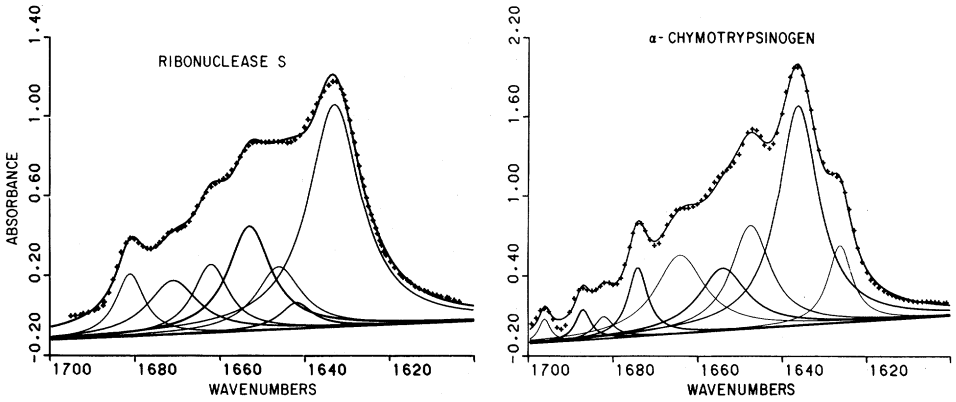

Fig. 2. Deconvolved FTIR spectra of ribonuclease S and α-chymotrypsinogen. (+ + + +) Experimental data; (———) calculated spectra. Curve-fitting and experimental conditions as in Fig. 1.

the $\%\beta$-structure values, $\%\beta$ (X-ray), estimated from X-ray data by Levitt and Greer [12], as shown in the next to last column of Table I. Major discrepancies only appear for ribonuclease A and α-chymotrypsin. Both of these proteins have the dominant β-band near 1637 cm^{-1} where α-helical proteins also have a band (see ref. 5 and Fig. 3). We have attempted to correct for this complication in the following manner. For α-chymotrypsinogen, where $\%A$ is in good agreement with reported X-ray data, we note that the ratio of the area of the more intense, low frequency β-band near 1637 cm^{-1} to the area of the β-band about 1675 cm^{-1} is 5.57 (i.e., $A_{1637} = 5.57 \cdot A_{1675}$). The assumption of the same constant intensity ratio seems reasonable for these two β-bands in other proteins when their frequencies are within ± 1 cm^{-1} of

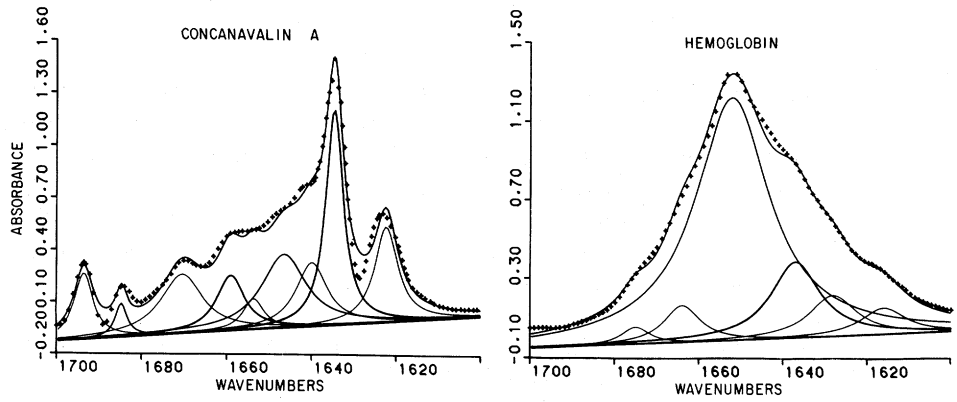

Fig. 3. Deconvolved FTIR spectra of ConA and hemoglobin. (+ + + +) Experimental data; (———) calculated spectra. Curve-fitting and experimental conditions as in Fig. 1. (Note evidence of over-deconvolution in the spectrum of Con A at 1630 cm^{-1}; this results from some of the amide I′ components having unusually narrow band widths.)

TABLE 1

AMIDE I′ COMPONENTS ASSOCIATED WITH β-SEGMENTS AS RELATED TO THE β-CONTENT OF PROTEINS [a]

Protein	cm^{-1}	A_i	A_{tot}	$\%A$	$\%\beta$		
					(FTIR)	(X-ray) [b]	(circular dichroism) [c]
Ribonuclease A	1630	10.24	74.70	55	50	46	44
	1637	26.42					36
	1676	4.13					
α-Chymotrypsin	1628	6.68	62.84	58	49	49	29
	1637	26.10					53
	1674	3.70					
α-Chymotrypsinogen	1627	8.50	82.07	47	47	45	36 [d]
	1636	25.44					
	1674	4.57					
Ribonuclease S	1633	26.29	60.32	54	54	53	37
	1671	6.09					33
Con A	1622	5.67	45.45	60	60	60	41
	1634	9.80					46
	1639	5.23					
	1670	6.55					

[a] By definition [3,4,6], the area of a Lorentzian band, $A = \pi \cdot s \cdot H$, where H is the peak height and s is the half width at half height; A_i is the area of the individual β-bands; A_{tot} is the sum of all the amide I′ component bands obtained by the curve-fitting procedure; $\%A = (\Sigma A_i / A_{tot}) \cdot 100$; $\%\beta$ (FTIR) is the corrected value of $\%A$ (see eqn. 1 in text).

[b] $\%\beta$(X-ray) is the percent β-structure determined by the method of Levitt and Greer [12].

[c] For $\%\beta$ (circular dichroism), the upper value is from the work of Provencher and Glockner [13], the lower value from the work of Chang et al. [14], except for α-chymotrypsinogen.

[d] Value obtained by Chen et al. (15).

the values quoted above. An alternative estimate of the β-content for ribonuclease A and α-chymotrypsin based on the corrected sum of the areas of the β-bands, can then be given as

$$\%\beta(\text{FTIR}) = \left[(A_{1675} + 5.57 \cdot A_{1675} + A_{1628}) / A_{tot} \right] \cdot 100$$

Column 5 of Table 1 gives these corrected values. The agreement with $\%\beta$ (X-ray) is markedly improved. (The correction is not applied to ribonuclease S or to Con A, because the observed frequencies are different and no correction is needed.)

The last column of Table 1 lists some literature values obtained by circular dichroism spectroscopy for comparison [13–15]. It is interesting to note that the β-content values obtained by FTIR spectroscopy by the described method are, generally, in better agreement with the values obtained from X-ray data by Levitt and Greer [12], than with the values obtained by several authors via circular dichroism.

For globular proteins it is, in general, not easy to determine objective values for the β-content or α-content because, among other complications, the endpoints of the

segments cannot always be clearly defined. A thoughtful discussion of these problems is given by Levitt and Greer [12]. Any reported values are somewhat subjective and depend on the applied definitions as well as the applied technology. The agreement between computed values based on precise X-ray data on one hand, and purely empirical values obtained by FTIR on the other, is therefore both satisfying and surprising.

Simplified description of the method and its accomplishments

The β-structure content of globular proteins can be estimated by combining: (a) Fourier deconvolution of high-quality infrared spectra in the 1600–1700 cm^{-1} region, (b) band assignments based on previous work and internal consistency and, (c) curve fitting by Gauss–Newton iteration. The obtained values are in good agreement with values calculated from X-ray data by the procedure of Levitt and Greer [12].

Acknowledgements

The authors thank Janine Brouillette for her invaluable assistance in preparing many of the solutions of the proteins and obtaining most of the FTIR spectra. Dr. William Damert and the staff of the ERRC computer center developed the Gauss–Newton iterative curve-fitting program ABCUS.

References

1 Griffiths, P.R. (1983) Science 222, 297–302.
2 Durig, J.R., ed. (1980) Analytical Applications of FT-IR to Molecular and Biological Systems, D. Reidel Publishing Co., Dordrecht, Holland.
3 Kauppinen, J.K., Moffat, D.J., Mantsch, H.H. and Cameron, D.G. (1981) Anal. Chem. 53, 1454–1457.
4 Maddams, W.F. and Southon, M.J. (1982) Spectrochim. Acta 38A, 459–466
5 Susi, H. and Byler, D.M. (1983) Biochem. Biophys. Res. Commun. 115, 391–397
6 Kauppinen, J.K., Moffat, D.J., Mantsch, H.H. and Cameron, D.G. (1981) Appl. Spectrosc. 35, 271–276.
7 Yang, W.-J. and Griffiths, P.R. (1981) Proc. Soc. Photo-Opt. Instrum. Eng. 289, 263–264
8 Purcell, J.M. and Susi, H. (1984) J. Biochem. Biophys. Methods 9, 193–199
9 Yang, W.J., Griffiths, P.R., Byler, D.M. and Susi, H. (1985) Appl. Spectrosc. 39, 282–287
10 Susi, H. (1969) Structure and Stability of Biological Macromolecules (Timasheff, S.N. and Fasman, G.D., eds.), pp. 575–663, Marcel Dekker, New York
11 Fukushima, K., Ideguchi, Y. and Miyazawa, T. (1963) Bull. Chem. Soc. Jpn. 36, 1301–1307.
12 Levitt, M. and Greer, J. (1977) J. Mol. Biol. 114, 181–293
13 Provencher, S.W. and Glockner, J. (1981) Biochemistry 20, 33–37
14 Chang, C.T., Wu, C.-S.C. and Yang, J.T. (1978) Anal. Biochem. 91, 13–31
15 Chen, Y.H., Yang, J.T. and Martinez, H.M. (1972) Biochemistry 11, 4120–4131